

I claim:

1. An arrangement including a plurality of interconnected cache servers, where at least some of the cache servers are connected to a data network, the arrangement

5 comprising:

Sub
a1
7
10) a cache selection module in each of said cache servers that, in response to a request for information, makes a determination, without consultation with other cache servers of said plurality of cache servers at the time said request arrives at said module, whether an attempt to service said request should be made a) at a specific one of said cache servers belonging to a set that includes said cache server of said module, or b) outside said arrangement.

2. The arrangement of claim 1 where said cache selection module makes its determination as to whether said request might be served by a specific one of said cache
15 servers by first determining whether said request for information might be served by the cache server of said module, and when a determination is made that said request cannot be served by the cache server of said module, then by determining whether said request might be served by a specific other one of said plurality of cache servers.

20 3. The arrangement of claim 1 where said cache selection module makes its determination as to whether said request might be served by a specific one of said cache servers by determining whether said request for information might be served by the cache server of said module, or by a specific one of said plurality of cache servers.

25 4. The claim 1 where said set includes a pre-designated one of said cache servers as a default cache server, and said cache selection module makes its determination as to whether said request might be served by a specific one of said cache servers by determining whether said request for information might be served by the cache server of said module, then, if necessary, whether said request might be served by a specific one of
30 said plurality of cache servers, then, if necessary, whether said request might be served by said default cache server.

5. The arrangement of claim 1 where said cache selection module makes its determination by consulting a table.

5 6. The arrangement of claim 5 where said table associates site specifications with the cache servers in said plurality of cache servers.

7. The arrangement of claim 5 where said table associates site or sub-site specifications, inclusively, with the cache servers in said plurality of cache servers.

10 8. The arrangement of claim 5 where said table effectively includes an entry for specifying a default cache server.

15 9. The arrangement of claim 1 where said cache selection module makes its determination by analyzing said request for information.

10. The arrangement of claim 9 said analyzing performs a transformation of said request for information.

20 11. The arrangement of claim 10 wherein said transformation is performed by means of a hash function applied to said request for information.

12. The arrangement of claim 1 where said cache selection module makes its determination by analyzing a site, or sub-site, specification contained in said request.

25 13. The arrangement of claims 12 said analyzing performs a transformation of said site, or sub-site, specification.

30 14. The arrangement of claim 13 wherein said transformation is performed by means of a hash function applied to said site, or sub-site, specification.

15. The arrangement of claim 1 further comprising an information-caching module in each of said cache servers, where said information-caching module decides whether to cache information of a given site in its cache server and operates accordance with an algorithm that comports with said determination made by said cache selection
5 module, achieving thereby a conformance that a given cache which is determined by said cache selection module to be the cache server that might service a request for information from some specified site, is also the cache server selected by said information-caching module for caching information from said some specified site.

10 16. The apparatus of claim 15 where said information-caching module selects a cache server for caching information from a particular site, or sub-site, based on the address of said site.

15 17. The apparatus of claim 15 where said information-caching module selects a cache server for caching information from a particular site, or sub-site, based proximity of the selected cache server to a community of users.

20 18. The apparatus of claim 15 where said information-caching module selects a cache server for caching information from a particular site, or sub-site, based on physical locations of said cache servers.

25 19. The apparatus of claim 15 where said information-caching module selects a cache server for caching information from a particular site, or sub-site, based a prearranged algorithm.

30 20. The apparatus of claim 15 where said information-caching module selects a cache server for caching information from a particular site, or sub-site, based on spare capacity of the selected cache server compared to spare capacity of unselected cache servers of said plurality of cache servers.

21. The apparatus of claim 15 where said information-caching module selects its own cache server for caching information.

22. The apparatus of claim 15 where said information-caching module selects a
5 default cache server for caching information.

23. The arrangement of claim 1, at least some of said cache servers further comprising a local cache for servicing requests for information received directly from said data network.

10

24. The arrangement of claim 1 where at least some of said cache servers further comprise a local cache which is accessed, in response to a request for information that arrives directly from said data network, before said cache selection module is accessed.

15 25. In an arrangement including a plurality of cache servers interconnected to form a virtual cache, a method for retrieving information from said virtual cache comprising the steps of:

receiving at one of said cache servers a request for information which specifies a site, or a sub-site, address that designates the source of said information;

20 converting said address destination to a designation that identifies a cache server in said virtual cache; and

directing said request for information to the identified cache server.

26. The method of claim 25 where said step of converting comprises obtaining a
25 cache server designation from a table.

27. The method of claim 25 where said step of converting comprises performing a transformation of said site, or sub-site, address to obtain a cache server designation.

28. In an arrangement including a plurality of cache servers interconnected to form a virtual cache, a method for caching information in said virtual cache, comprising the steps of:

receiving a request for information which specifies a site, or a sub-site, address
5 that designates the source of said information;
evaluating whether said request can be serviced by said virtual cache,
when said step of evaluating determines that said virtual cache cannot service said
request, routing said request to said site, or sub-site address,
receiving information responsive to said request for information; and
10 based on address granularity no finer than sub-site address, assigning one of said
cache servers to cache said information.

29. The method of claim 28 where said assigning is based on a transformation of said site, or sub-site, address.

30. The method of claim 28 where said assigning is to a specific one of said cache servers that is designated a default cache server.

31. The method of claim 28 where said assigning is based on spare capacity that is available at the cache servers.

32. The method of claim 28 where said assigning is based physical locations of said cache servers.